

Optical Diffraction in Close Proximity to Plane Apertures. IV. Test of a Pseudo-Vectorial Theory

Volume 111

Number 1

January-February 2006

Klaus D. Mielenz

National Institute of Standards
and Technology,
Gaithersburg, MD 20899-8440

klausm@hereintown.net

Rayleigh's pseudo-vectorial theory of the diffraction of polarized light by apertures which are small compared to the wavelength of light is analyzed with respect to its mathematical rigor and physical significance. It is found that the results published by Rayleigh and Bouwkamp for *s*-polarized incident do not obey the conditions assumed in their derivation and must therefore be dismissed. It is also found that the theory leads to paradoxical predictions concerning the polarization of the diffracted field, so that the pseudo-vectorial approach is intrinsically incapable of describing polarization effects.

Key words: diffraction; narrow slits; optics; polarization; pseudo-vectorial theory; small circular apertures; scalar theory; transmission coefficients.

Accepted: January 4, 2006

Available online: <http://www.nist.gov/jres>

1. Introduction

Most applications of diffraction theory in optics are based on the scalar approximation of Huygens' principle, and it is usually understood that polarization is ignored. A notable exception to this practice is the classical pseudo-vectorial theory formulated by Rayleigh [1,2] and Bouwkamp [3] for the diffraction of plane-polarized light by small metallic apertures as illustrated in Fig. 1. The results obtained by these authors implied large polarization effects in the case of aperture sizes smaller than the wavelength of light, λ . Specifically, Rayleigh and Bouwkamp predicted that unpolarized light diffracted by very narrow aperture will become strongly *p*-polarized, as indicated by the ratio $\tau_{RB}^{(p)}/\tau_{RB}^{(s)}$ of the transmission coefficients of circular apertures and slits illuminated by normally incident *p*- or *s*-polarized light plotted in Fig. 2 for the range $0 < kw \leq 0.75$, where $2w$ is the aperture width and $k = 2\pi/\lambda$ is the circular wave number. Regarding the case of slits,

Rayleigh commented: "It appeared that if the width of the slit is very small in comparison with the wavelength, there is a much more free passage when the electric vector is perpendicular to the slit than when it is parallel to the slit, so that unpolarized light incident upon the screen will, after passage, appear polarized in the former manner." While this may be plausible as far as slits are concerned, it seems impossible that a circular aperture can behave in the same manner because normal incidence is assumed so that symmetry would dictate a complete absence of polarization. Apparently Rayleigh had similar doubts. He remarked that the *p*-solution has "no simple application" in the case of circular apertures, but nonetheless the discrepancy is disturbing as the Rayleigh-Bouwkamp theory is commonly referred to as a "rigorous" theory, and in this connection it should also be noted that later authors, such as Levine and Schwinger [4,5], considered only the *s*-solution and, in doing so, made no reference to polarization. In the present paper the Rayleigh-Bouwkamp

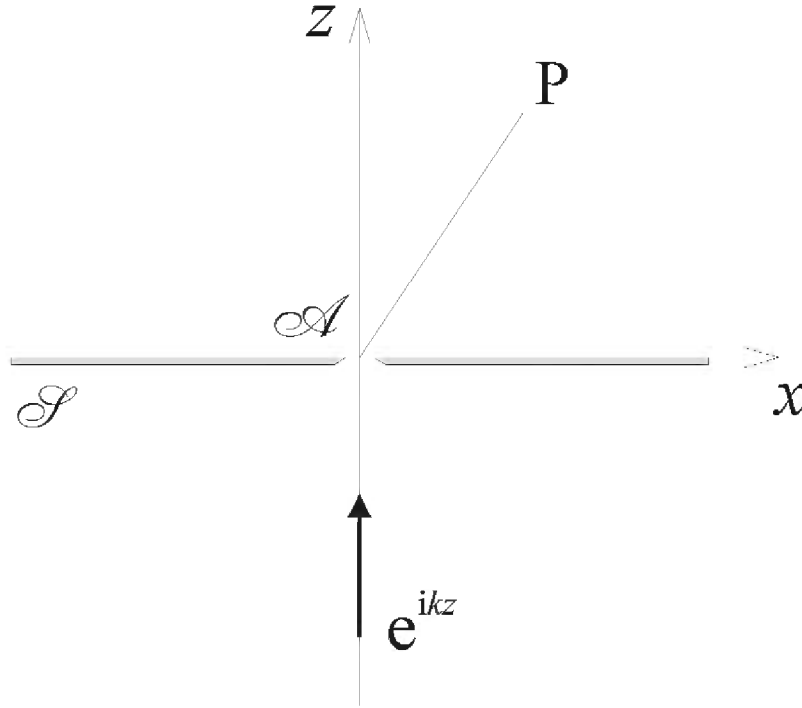


Fig. 1. The Rayleigh-Bouwkamp diffraction problem. A very narrow aperture \mathcal{A} in a metallic screen \mathcal{S} which is located in the xy -plane is illuminated by a monochromatic, p - or s -polarized plane wave e^{ikz} , and solutions are constructed so that the optical field is everywhere continuous with continuous first derivatives.

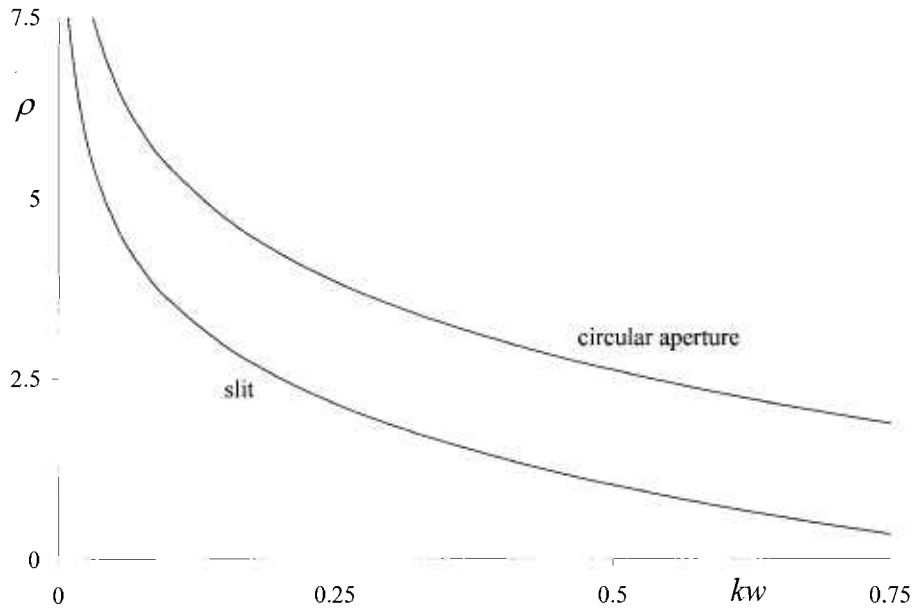


Fig. 2. Polarization ratios, $\rho = \tau_{\text{RB}}^{(p)} / \tau_{\text{RB}}^{(s)}$, of circular apertures and slits vs. aperture size kw according to Bouwkamp [3, p.71 ff].

theory will be reexamined with respect to its physical significance, mathematical rigor, and accuracy.

Rayleigh considered a monochromatic plane wave of unit irradiance, $u_0 = e^{ikz}$, which is incident from the half space $z < 0$ on a circular aperture or slit \mathcal{A} contained in an infinite, infinitesimally thin, opaque screen \mathcal{S} that lies in the xy -plane of a cartesian coordinate system (Fig. 1). He also assumed that the diffracted field in the half space $z \geq 0$ obeys the Rayleigh-Sommerfeld integral equations,

$$u^{(p)}(P) = -\frac{1}{2\pi} \int_{\mathcal{A}} dQ \frac{\partial u^{(p)}(Q)}{\partial z} \frac{e^{ikQP}}{QP}, \quad z \geq 0, \quad (1a)$$

$$u^{(s)}(P) = -\frac{1}{2\pi} \int_{\mathcal{A}} dQ u^{(s)}(Q) \frac{\partial}{\partial z} \left(\frac{e^{ikQP}}{QP} \right), \quad z \geq 0, \quad (1b)$$

where $Q = (\xi, \eta, 0)$ is a point in \mathcal{A} , $P = (x, y, z)$ is an arbitrary point in space, QP is the distance between them, $u^{(p,s)}$ are scalar wave functions, and $\partial u^{(p)}(Q)/\partial z$ and $u^{(s)}(Q)$ are zero on the opaque portion of \mathcal{S} as if the screen is a perfect metallic reflector and the incident light is p - or s -polarized. Furthermore, Rayleigh assumed that field on the source side of \mathcal{S} is a modified geometrical field of the form

$$u^{(p,s)}(P) = e^{ikz} \pm e^{-ikz} + u_{\pm}^{(p,s)}(P), \quad z \leq 0, \quad (1c)$$

and hence the objective of the theory is to find the unknown quantities $u_{\pm}^{(p,s)}(P)$ and $u_{\pm}^{(p,s)}(P)$ so that the overall field is continuously differentiable in the aperture plane.

From Eq. (1c) it follows that $\partial u^{(p)}/\partial z$ and $u^{(s)}$ will be continuous for $z = 0$ if

$$\frac{\partial u^{(p)}(x, y, 0)}{\partial z} = \frac{\partial u_{-}^{(p)}(x, y, 0)}{\partial z}, \quad u^{(s)}(x, y, 0) = u_{-}^{(s)}(x, y, 0), \quad (2a)$$

so that

$$u_{-}^{(p,s)}(x, y, z) = \mp u^{(p,s)}(x, y, -z), \quad z \geq 0, \quad (2b)$$

$$u_{-}^{(p,s)}(x, y, z) = e^{ikz} \pm e^{-ikz} \mp u^{(p,s)}(x, y, -z), \quad z \leq 0. \quad (2c)$$

Thus, $u^{(p)} = 2 - u_{-}^{(p)}$, $\partial u^{(s)}/\partial z = 2ik - \partial u_{-}^{(s)}/\partial z$ when $z = 0$, so that $u^{(p)}$ and $\partial u^{(s)}/\partial z$ will also be continuous if

$$u^{(p)}(x, y, 0) = 1, \quad \frac{1}{ik} \frac{\partial u^{(s)}(x, y, 0)}{\partial z} = 1 \quad (2d)$$

The conditions [Eq. (2d)] represent the underlying integral equations of the Rayleigh-Bouwkamp and other so-called “rigorous” diffraction theories which are intended to solve the wave equation in the presence of prescribed boundary conditions. To this author’s knowledge, all existing solutions with the sole exception of Sommerfeld’s half-plane theory are approximations of some sort and in the cases considered here these approximations take the form of ascending power series in kw which were derived on the assumption that the unknown aperture distributions in Eqs. (1a,b) can be expressed in the form

$$\begin{aligned} \frac{1}{ik} \frac{\partial u^{(p)}(Q)}{\partial z} &= -\frac{i}{kw} \left[\frac{p_0}{W} + (kw)^2 p_1 W + \dots \right], \\ u^{(s)}(Q) &= -ikw [s_1 W + (kw)^2 s_2 W^3 + \dots], \end{aligned} \quad (3)$$

where $W = \sqrt{1 - (q/w)^2}$, q is the distance of Q from the coordinate origin, and the coefficients p_n and s_n are to be chosen so that Eqs. (2d) are satisfied. The determination of these coefficients is tedious and will not be repeated here as the results have been detailed in Ref. [3]. In the following the corresponding solutions will be denoted by the subscript ‘RB’ and it will be sufficient to consider only the initial terms of the expansions [Eq. (3)] by assuming, as Rayleigh did originally, that $kw \rightarrow 0$.

2. Circular Apertures

For circular apertures it is appropriate to use polar coordinates inside the aperture and in the plane of observation, so that

$$P = (\rho \cos \varphi, \rho \sin \varphi, z), \quad \rho^2 = x^2 + y^2, \quad (4a)$$

$$Q = (q \cos \chi, q \sin \chi, 0), \quad q^2 = \xi^2 + \eta^2, \quad (4b)$$

$$dQ = q dq d\chi, \quad QP^2 = z^2 + \rho^2 + q^2 - 2\rho q \cos(\varphi - \chi). \quad (4c)$$

The initial terms of the trial solutions of Eq. (3) were determined by Rayleigh and Bouwkamp as $p_0 = s_1 = -2/\pi$ and thus the first-order solutions for circular apertures are

$$u_{\text{RB}}^{(p)}(P) = \frac{1}{\pi^2} \int_0^w dq \int_0^{2\pi} d\chi \frac{q e^{ikQP}}{\sqrt{w^2 - q^2} QP}, \quad kw \rightarrow 0, \quad z \geq 0, \quad (4d)$$

$$u_{\text{RB}}^{(s)}(P) = -\frac{ik}{\pi^2} \int_0^w dq \int_0^{2\delta} d\chi q \sqrt{w^2 - q^2} \frac{\partial}{\partial z} \left(\frac{e^{ikQP}}{QP} \right) = -\frac{k^2 z}{\pi^2} \int_0^w dq \int_0^{2\delta} d\chi q \sqrt{w^2 - q^2} e^{ikQP} \left(\frac{1}{QP^2} + \frac{i}{kQP^3} \right), \quad kw \rightarrow 0, \quad z \geq 0. \quad (4e)$$

2.1 Fresnel Approximation

The evaluation of Eqs. (4d,e) in the Fresnel limit is simple. When $z \gg \lambda$, and thus $z \gg q$, the dependence of QP on the aperture coordinates can be ignored, so that

$$QP \sim \sqrt{z^2 + \rho^2} = r, \quad \frac{z}{r} = \cos \theta \quad (5a)$$

and

$$u_{\text{RB}}^{(p)}(P) = -\frac{w}{\pi^2} \frac{e^{ikr}}{r} \int_0^1 d(q/w) (q/w)^{2\pi} \int_0^{2\pi} d\chi = \frac{2w e^{ikr}}{\pi r}, \quad z \gg \lambda, \quad (5b)$$

$$u_{\text{RB}}^{(s)}(P) = \frac{2k^2 w^3 \cos \theta e^{ikr}}{\pi r} \int_0^1 d(q/w) (q/w) \sqrt{1 - (q/w)^2}, \\ = \frac{2k^2 w^3 \cos \theta e^{ikr}}{3\pi r}, \quad z \gg \lambda \quad (5c)$$

where $q/w = \sin \alpha$ was substituted in both integrals. Hence, the total radiant fluxes transmitted by the aperture and incident on the inside of a large hemisphere with radius r are given by

$$\Phi^{(p)} = 2\pi r^2 \int_0^{\pi/2} d\theta \sin \theta |u_{\text{RB}}^{(p)}|^2 = \frac{8w^2}{\pi}, \quad (5d)$$

$$\Phi^{(s)} = 2\pi r^2 \int_0^{\pi/2} d\theta \sin \theta |u_{\text{RB}}^{(s)}|^2 = \frac{8k^4 w^6}{9\pi} \int_0^{\pi/2} d\theta \sin \theta \cos^2 \theta \\ = \frac{8k^4 w^6}{27\pi}, \quad (5e)$$

and the corresponding transmission coefficients of the aperture are

$$\tau_{\text{RB}}^{(p)} = \frac{\Phi^{(p)}}{\Phi_0} = \frac{8}{\pi^2}, \quad \tau_{\text{RB}}^{(s)} = \frac{\Phi^{(s)}}{\Phi_0} = \frac{8(kw)^4}{27\pi^2}, \quad kw \rightarrow 0, \quad (5f)$$

where $\Phi_0 = \pi w^2$ is the geometrical flux incident on the aperture.

These results can now be checked for consistency with Eqs. (2d) and (3) by using an alternative expression [3,6] for transmission coefficients,

$$\tau' = \frac{1}{A} \int_A dQ |u(Q)|^2, \quad (6a)$$

where A is the area of the aperture and $u(Q)$ is the field inside it. According to Eq. (2d) we expect $u_{\text{RB}}^{(p)}(Q) = 1$ and therefore

$$(\tau_{\text{RB}}^{(p)})' = 1, \quad (6b)$$

whereas from Eq. (3) and $s_1 = -2/\pi$ it follows that

$$(\tau_{\text{RB}}^{(s)})' = \frac{(s_1 kw)^2}{\pi w^2} \int_0^1 d\chi \int_0^{2\pi} d\chi q [1 - (q/w)^2] = \frac{2(kw)^2}{\pi}, \quad kw \rightarrow 0. \quad (6c)$$

The p -coefficients of Eqs. (5f) and (6b) agree reasonably well. On the other hand, there is a considerable difference between the corresponding s -coefficients.

2.2 Near-Field Solution

In the immediate proximity of the aperture plane the exponential e^{ikQP} is effectively constant and equal to 1 inside the aperture, and Eqs. (4d,e) can be simplified as follows,

$$u_{\text{RB}}^{(p)}(P) = \frac{1}{\pi^2} \int_0^w \frac{q dq}{\sqrt{w^2 - q^2}} \int_0^{2\pi} \frac{d\chi}{QP}, \quad z \ll \lambda \quad (7a)$$

$$u_{\text{RB}}^{(s)}(P) = -\frac{k^2 z}{\pi^2} \int_0^w dq q \sqrt{w^2 - q^2} \int_0^{2\pi} d\chi \left(\frac{1}{QP^2} + \frac{i}{kQP^3} \right), \quad z \ll \lambda. \quad (7b)$$

As these wave functions must be rotationally symmetrical about the z -axis it suffices to evaluate the χ -integrals appearing in these expressions for $\varphi = 0$. Substituting $\alpha = (\chi - \pi)/2$, we find

$$QP^2 = z^2 + \rho^2 + q^2 - 2\rho q \cos \chi = \frac{4\rho q}{m} (1 - m \sin^2 \alpha), \quad (8a)$$

$$m = \frac{4\rho q}{z^2 + (\rho + q)^2}, \quad (8b)$$

and hence

$$\int_0^{2\pi} \frac{d\chi}{QP} = 2 \sqrt{\frac{m}{\rho q}} \int_0^{\pi/2} \frac{d\alpha}{(1 - m \sin^2 \alpha)^{1/2}} = 2 \sqrt{\frac{m}{\rho q}} K(m), \quad (8c)$$

$$\int_0^{2\pi} \frac{d\chi}{QP^2} = \frac{m}{\rho q} \int_0^{\pi/2} \frac{d\alpha}{1 - m \sin^2 \alpha} = \frac{\pi m}{2\rho q \sqrt{1 - m}}, \quad (8d)$$

$$\int_0^{2\pi} \frac{d\chi}{Q P^3} = \frac{1}{2} \left(\frac{m}{\rho q} \right)^{3/2} \int_0^{\pi/2} \frac{d\alpha}{(1 - m \sin^2 \alpha)^{3/2}} = \frac{1}{2} \left(\frac{m}{\rho q} \right)^{3/2} \frac{E(m)}{1 - m}, \quad (8c)$$

where $K(m)$ and $E(m)$ are the complete elliptical integrals of the first and second kinds [7,8].

Upon substitution of Eq. (8c), Eq. (7a) is reduced to the single integral

$$u_{RB}^{(p)}(P) = \frac{2}{\pi^2 \sqrt{\rho/w}} \int_0^1 \frac{d(q/w) \sqrt{mq/w} K(m)}{\sqrt{1 - (q/w)^2}}, \quad z \ll \lambda, \quad (9a)$$

where relative coordinates are used. This expression was evaluated by numerical integration to analyze its behavior in the aperture plane, using a polynomial approximation for $K(m)$ given in Ref. [7] and assuming $z/w = 0$ as well as $z/w = 0.01$ in order to find $\partial u_{RB}^{(p)}/\partial z$ by numerical differentiation. The results of these computations are listed in the left-hand portion of Table 1, which shows that they are in satisfactory agreement with the values expected from Eqs. (2d) and (4d).

The corresponding expressions for $u_{RB}^{(s)}$ are obtained from Eqs. (7b) and (8d) and can be expressed as follows,

$$u_{RB}^{(s)}(P) = kz \left[\frac{k w A}{\rho/w} + \frac{i B}{(\rho/w)^{3/2}} \right], \quad \frac{1}{ik} \frac{\partial u_{RB}^{(s)}(P)}{\partial z} = \frac{B}{(\rho/w)^{3/2}} - \frac{i A}{\rho/w}, \quad (10a)$$

where

$$A = -\frac{1}{2\pi} \int_0^1 \frac{d(q/w) m \sqrt{1 - (q/w)^2}}{\sqrt{1 - m}}, \quad z \ll \lambda, \quad (10b)$$

$$B = -\frac{1}{2\pi^2} \int_0^1 \frac{d(q/w) m^{3/2} \sqrt{1 - (q/w)^2} E(m)}{(1 - m) \sqrt{q/w}}, \quad z \ll \lambda, \quad (10c)$$

and $E(m)$ can also be expressed by a polynomial approximation [7]. It was ascertained that the integrals A and B are everywhere finite, and hence it follows from the first Eq. (10a) that the aperture values of $u_{RB}^{(s)}$ are identically equal to zero. This is in disagreement with Eq. (4c), and it was also found that the computed aperture values of $(1/ik) \partial u_{RB}^{(s)}/\partial z$ do not satisfy the second integral equation, [Eq. (2d)]. These discrepancies are documented in the right-hand portion of Table 1.

3. Slits

The second case considered by Rayleigh and Bouwkamp is that of an infinitely long, narrow slit of width $2w$ which is centered on the y -axis of a rectangular coordinate system. As the resulting diffraction pattern will be independent of y , it is sufficient to compute its variation in the xz -plane by assuming

$$P = (x, 0, z), \quad Q = (\xi, \eta, 0), \quad (11a)$$

and then Eqs. (1a,b) are reduced to

$$u_{RB}^{(p)}(P) = -\frac{1}{2\pi} \int_{-w}^w d\xi \frac{\partial u_{RB}^{(p)}(\xi, 0)}{\partial z} \int_{-\infty}^{\infty} d\eta \frac{e^{ik\sqrt{(\xi-x)^2 + \eta^2 + z^2}}}{\sqrt{(\xi-x)^2 + \eta^2 + z^2}} = -\frac{i}{2} \int_{-w}^w d\xi \frac{\partial u_{RB}^{(p)}(\xi, 0)}{\partial z} H_0^{(1)}(\beta), \quad \beta = k\sqrt{(\xi-x)^2 + z^2}, \quad (11b)$$

$$u_{RB}^{(s)}(P) = -\frac{i}{2} \int_{-w}^w d\xi u_{RB}^{(s)}(\xi, 0) \frac{\partial H_0^{(1)}(\beta)}{\partial z} = -\frac{ik^2 z}{2} \int_{-w}^w d\xi u_{RB}^{(s)}(\xi, 0) \frac{H_1^{(1)}(\beta)}{\beta}, \quad (11c)$$

where the infinite η -integral in Eq. (10b) was evaluated in terms of the Hankel function $H_0^{(1)}(\beta)$ [8] and the relationship $dH_0^{(1)}(\beta)/d\beta = -H_1^{(1)}(\beta)$ was used in Eq. (10c). According to Rayleigh and Bouwkamp the initial terms of the series expansions [Eq. (3)] are $p_0 = 1/(o - i\pi/2)$ and $s_1 = 1$ where $o = \ln(\gamma kw/4)$ and $\gamma = 0.577..$

Table 1. Aperture values of wave functions and normal derivatives for circular apertures of width $kw = 0.1$, as computed from Eqs. (9a) and (10a,b,c) and expected from Eqs. (2d) and (3)

ρ/w	$u_{RB}^{(p)}(Q)$		$(1/ik) \partial u_{RB}^{(p)}(Q)/\partial z$		$u_{RB}^{(s)}(Q)$		$(1/ik) \partial u_{RB}^{(s)}(Q)/\partial z$	
	computed	expected	computed	expected	computed	expected	computed	expected
0.01	0.99	1	-0.62	-0.66	0	0.064 i	0.5 i	1
0.25	0.99	1	-0.64	-0.66	0	0.062 i	-30 + 0.5 i	1
0.50	0.99	1	-0.72	-0.74	0	0.055 i	-216 + 0.4 i	1
0.75	0.99	1	-0.94	-0.96	0	0.042 i	-558 + 0.3 i	1
0.99	0.98	1	-3.91	-4.51	0	0.009 i	-276	1

is Euler's constant, and hence the first-order solutions for slits are

$$u_{RB}^{(p)}(P) = -\frac{ip_0}{2w} \int_{-w}^w \frac{d\xi H_0^{(1)}(\beta)}{\sqrt{1-(\xi/w)^2}}, \quad kw \rightarrow 0, \quad z \geq 0, \quad (11d)$$

$$u_{RB}^{(s)}(P) = -\frac{k^3 wz}{2} \int_{-w}^w \frac{d\xi \sqrt{1-(\xi/w)^2} H_1^{(1)}(\beta)}{\beta}, \quad kw \rightarrow 0, \quad z \geq 0. \quad (11e)$$

3.1 Fresnel Approximation

At large distances from the slit ($z \gg w$), the phase difference β defined in Eq. (11b) is effectively a large constant and equal to $kr = k\sqrt{x^2 + z^2}$, and the Hankel functions in Eqs. (11d,e) can be replaced by their asymptotic expansions

$$H_0^{(1)}(\beta) \sim \sqrt{\frac{2}{\pi kr}} e^{i(kr-\pi/4)}, \quad \frac{H_0^{(1)}(\beta)}{\beta} \sim \sqrt{\frac{2}{\pi(kr)^3}} e^{i(kr-3\pi/4)}, \quad (12a)$$

so that

$$\begin{aligned} u_{RB}^{(p)}(P) &= -\frac{i\sqrt{2}p_0 e^{i(kr-\pi/4)}}{\sqrt{\pi kr}} \int_0^1 \frac{d(\xi/w)}{\sqrt{1-(\xi/w)^2}} \\ &= -\frac{i\sqrt{\pi} e^{i(kr-\pi/4)}}{(o-i\pi/2)\sqrt{2kr}}, \quad z \gg \lambda, \end{aligned} \quad (12b)$$

$$\begin{aligned} u_{RB}^{(s)}(P) &= -\frac{\sqrt{2}(kw)^2 \cos\theta e^{i(kr-3\pi/4)}}{\sqrt{\pi kr}} \int_0^1 d(\xi/w) \sqrt{1-(\xi/w)^2} \\ &= -\frac{\sqrt{\pi}(kw)^2 \cos\theta e^{i(kr-3\pi/4)}}{\sqrt{8kr}}, \quad z \gg \lambda, \end{aligned} \quad (12c)$$

where $\cos\theta = z/r$ and the two integrals were again evaluated by substituting $\xi/w = \sin\alpha$. These expressions can now be used to compute the quantities

$$\Delta\Phi^{(p,s)} = r\Delta y \int_{-\pi/2}^{\pi/2} d\theta |u_{RB}^{(p,s)}|^2, \quad (12d)$$

which represent the radiant flux emerging from a slit segment of length Δy and falling on the corresponding annular segment of a semicylinder of radius r which is centered on the y -axis. As the geometrical flux incident on the slit segment is $\Delta\Phi_0 = 2w\Delta y$, this leads at once to the following expressions for the transmission coefficients of the slit,

$$\begin{aligned} \tau_{RB}^{(p)} &= \frac{\Delta\Phi^{(p)}}{\Delta\Phi_0} = \frac{\pi^2 |p_0|^2}{4kw} = \frac{\pi^2 / 4kw}{o^2 + \pi^2 / 4}, \quad \tau_{RB}^{(s)} = \frac{\Delta\Phi^{(s)}}{\Delta\Phi_0} \\ &= \frac{\pi^2 (kw)^3}{32}, \quad kw \rightarrow 0. \end{aligned} \quad (12e)$$

As in Sec. 2.1 alternative expressions for these transmission coefficients can be obtained from Eq. (6a), which in this case is equivalent to

$$\tau' = \frac{1}{w} \int_0^w d\xi |u(\xi)|^2. \quad (13a)$$

Hence we find, using Eqs. (2d) and (3) with $s_1 = 1$,

$$(\tau_{RB}^{(p)})' = 1, \quad (13b)$$

$$(\tau_{RB}^{(s)})' = kw \int_0^w d\xi [1 - (\xi/w)^2] = \frac{2(kw)^2}{3}, \quad kw \rightarrow 0. \quad (13c)$$

Both of these results disagree with the corresponding values in Eq. (12e).

3.2 Near-Field Solution

The integrals [Eq. (11d,e)] can be evaluated without simplifying assumptions for arbitrary values of z , and it was found that they are finite in the aperture plane in spite of the singularities of $H_0^{(1)}(\beta)$ and $H_1^{(1)}(\beta)/\beta$ for $\beta = 0$.

For p -polarized light we find

$$\begin{aligned} u_{RB}^{(p)}(P) &= \frac{1}{2M} \int_{-1}^1 \frac{d(\xi/w) \{ [\pi J_0(\beta)/2 + oY_0(\beta)] - i[oJ_0(\beta) - \pi Y_0(\beta)/2] \}}{\sqrt{1-(\xi/w)^2}}, \end{aligned} \quad (14)$$

where J_n and Y_n are Bessel functions of the first and second kinds and $M = o^2 + \pi^2/4$. The integral [Eq. (14)] was evaluated numerically for $kw = 0.1$ and $z = 0$ as well as $z = 0.01w$, yielding the results for $u_{RB}^{(p)}$ and $\partial u_{RB}^{(p)}/\partial z$ shown in the left-hand portion of Table 2. The table also shows the corresponding values expected from Eqs. (2d) and (3), and the agreement was judged satisfactory in view of the approximations made in Rayleigh's derivations.

The corresponding expression obtained from Eq. (11e) is

$$u_{RB}^{(s)}(P) = -\frac{k^2 wz}{2} (C + iD), \quad (15a)$$

Table 2. Aperture values of wave functions and normal derivatives for slits of width $kw = 0.1$, as computed from Eqs. (14) and (15a,b,c) and expected from Eqs. (2d) and (3)

ρ/w	$u_{RB}^{(p)}(Q)$		$(1/ik)\partial u_{RB}^{(p)}(Q)/\partial z$		$u_{RB}^{(s)}(Q)$		$(1/ik)\partial u_{RB}^{(s)}(Q)/\partial z$	
	computed	expected	computed	expected	computed	expected	computed	expected
0.01	0.76	1	0.75 + 2.0 i	0.76 + 2.1 i	0	-0.10 i	8.6 - 38 i	1
0.25	0.76	1	0.78 + 2.1 i	0.79 + 2.1 i	0	-0.10 i	8.3 - 37 i	1
0.50	0.76	1	0.87 + 2.3 i	0.89 + 2.4 i	0	-0.09 i	7.4 - 33 i	1
0.75	0.76	1	1.1 + 3.1 i	1.2 + 3.1 i	0	-0.05 i	5.8 - 25 i	1
0.99	0.75	1	4.7 + 13 i	5.5 + 15 i	0	-0.01 i	1.6 - 4 i	1

where

$$C = \int_{-1}^1 \frac{d(\xi/w) \sqrt{1 - (\xi/w)^2} J_1(\beta)}{\beta},$$

$$D = \int_{-1}^1 \frac{d(\xi/w) \sqrt{1 - (\xi/w)^2} Y_1(\beta)}{\beta}. \quad (15b)$$

It was found that the integrals [Eq. (15b)] are finite in the aperture plane, so that according to Eq. (15a) the aperture values of $u_{RB}^{(s)}$ are identically zero and thus do not agree with the assumed values of Eq. (3). Equation (15a) also shows that

$$\frac{1}{ik} \frac{\partial u_{RB}^{(s)}(P)}{\partial z} = \frac{kw}{2} (D - iC), \text{ if } z = 0, \quad (15c)$$

which obviously disagrees with Eq. (2d). The various numerical results obtained for $u_{RB}^{(s)}$ and $\partial u_{RB}^{(s)}/\partial z$ are listed on the right-hand side of Table 2.

4. Conclusion

Based on the above-mentioned results and illustrated by the numerical examples shown in Tables 1 and 2 there can be no doubt that the Rayleigh-Bouwkamp theory is mathematically flawed and fails to give a credible account of the diffraction of polarized light. The reasons for this conclusion are twofold.

The near-field solutions derived in Secs 2.2 and 3.2 for p -polarized incident light ($\partial u_0/\partial z = 0$ on \mathcal{S}) are in approximate agreement with the conditions in Eqs. (2d) and (3) that were assumed in their derivation, and thus appear to be free of contradictions. On the other hand, the corresponding solutions for s -polarization ($u_0 = 0$ on \mathcal{S}) do not obey these conditions at all. This discrepancy was found for circular apertures and slits alike, and thus points to a general problem associated with the second integral equation [Eq. (2d)] as such. Bouwkamp [3] mentioned that this integral equation [Eq. (2d)] is notoriously difficult as it requires differentiation with

respect to z under the integral sign and can lead to divergent solutions. He pointed out that these divergencies can be eliminated, but this was not attempted in this work. In any case, the polarization effects shown in Fig. 1 must be dismissed as wrong.

A second, and most likely uncorrectable, problem arises from the fact that the ‘polarization effects’ predicted by the Rayleigh-Bouwkamp theory run counter to common experience. As mentioned in the Introduction, a polarized field behind a circular aperture illuminated by normally incident, unpolarized light is impossible for reasons of symmetry. Yet the existence of such a field is implied by the upper curve in Fig. 1, and the discrepancy persists even when the suspect transmission coefficient $\tau_{RB}^{(s)}$ in Eq. (5f) is replaced by the improved coefficient $(\tau_{RB}^{(s)})'$ in Eq. (6c). Additionally, the different values obtained in Secs. 2.1 and 3.1 for the far- and near-field transmission coefficients seemed paradoxical because the polarization of light cannot change during free-space propagation. As this discrepancy was encountered for both states of polarization, it also cannot be blamed on the mathematical problem mentioned in the previous paragraph.

Hence it is concluded that the Rayleigh-Bouwkamp theory cannot be relied upon for describing the diffraction of polarized light by small apertures. Virtually the same determination was made in an earlier paper [6] on the feasibility of using the Rayleigh-Sommerfeld boundary-value integrals for analyzing diffraction by apertures of arbitrary size, and thus the overall conclusion is that the inability to account for polarization effects is an inherent property of scalar diffraction theories, no matter what boundary conditions are assumed. The fact of the matter appears to be that, in scalar diffraction theory, the boundary conditions $\partial u_0/\partial z = 0$ or $u_0 = 0$ are necessary and sufficient to satisfy Helmholtz’ theorem. As it happens, the same conditions also pertain to the metallic reflection of p - or s -polarized light but it does not follow that they are sufficient to ensure conformance with electromagnetic theory.

In Ref. [6] it was also shown that a modified theory of unpolarized diffraction can be formulated in which the Rayleigh-Sommerfeld integrals are used to describe a continuous, bidirectional flow of energy in the near zones on either side of an aperture. The transmission coefficients computed in this manner are similar to those published by Levine and Schwinger [4,5] for sub-wavelength aperture sizes, so that a special theory for this region may not be needed.

5. References

- [1] Lord Rayleigh, Phil. Mag. **43**, 259-272 (1897).
- [2] Lord Rayleigh, Proc. Roy. Soc. (A) **89**, 194-219 (1897).
- [3] C. J. Bouwkamp, Rep. Progr. Phys. (London) **17**, 35-100 (1953).
- [4] H. Levine and J. Schwinger, Phys. Rev. **74**, 958-974 (1948).
- [5] H. Levine and J. Schwinger, Phys. Rev. **75**, 1423-1432 (1949).
- [6] K. D. Mielenz, J. Res. Natl. Inst. Stand. Technol. **109**, 457-464 (2004).
- [7] M. Abramowitz and I. A. Stegun, Handbook of Mathematical Functions, U.S. Gov. Printing Office (1972).
- [8] I. S. Gradshteyn and I. M. Ryzhik, Table of Integrals, Series, and Products, Academic Press, New York etc. (1980).

About the author: Klaus D. Mielenz is a physicist and retired Chief of the Radiometric Physics Division of NIST Physics Laboratory. The National Institute of Standards and Technology is an agency of the Technology Administration, U.S. Department of Commerce.